PERCEPTUAL GROUPING BY PROXIMITY AND SIMILARITY IN THE PERIPHERAL VISION *

Jiang Yi, Han Shihui

(Department of Psychology, Center for Brain and Cognitive Sciences, Peking University, Beijing 100871)  
(Learning & Cognition Lab, Capital Normal University, Beijing 100089)

Abstract Our previous studies suggest that proximity dominates similarity in perceptual grouping in the foveal vision. The current work further investigated the difference between the two grouping processes in the peripheral vision using target detection and discrimination tasks in two experiments. Subjects were required to identify orientations of perceptual groups defined by either proximity or similarity that were presented in the left or the right visual fields. We found that reaction times were faster in the proximity than similarity grouping conditions regardless of whether subjects detected perceptual groups with specific orientations or discriminated orientations of perceptual groups in each trial. In addition, the dominance of proximity over similarity was not different between stimulus arrays presented in the left or the right visual fields. However, response speed did not differ between conditions when grouping was defined by similarity of orientation and closure. The results suggest that, in the peripheral vision, proximity grouping takes place earlier than grouping by similarity whereas similarity grouping by different geometrical features operates with the same speed.

Key words perceptual grouping; proximity; similarity; peripheral vision.

1 Introduction

Perceptual grouping refers to the function of the visual system to segment the visual field into perceptual groups or objects. It is widely accepted that perceptual grouping occurs at an early stage of the visual processing stream[1-4]. Gestalt psychologists suggest that proximity and similarity are two of the fundamental principles guiding perceptual grouping[5,6]. The principle of proximity states that nearby objects in the visual field tend to be perceived as a common group. The principle of similarity states that, all else being equal, elements that have similar features tend to be grouped together.

Several studies suggest that proximity plays a more critical role than similarity in guiding perceptual grouping. Ben – Av and Sagi[7] had subjects identify orientations of stimulus arrays composed of local elements and presented in the center of the visual field. They found that subjects responded faster to the orientations of perceptual groups formed by proximity than by similarity of shape. Han et al. [8-10] asked subjects to discriminate global shapes composed of local elements that were grouped by either proximity or similarity. They also found faster reaction times (RTs) to proximity relative to similarity stimuli in the foveal vision. In addition, Han et al. found that RTs to identify the perceptual groups were faster when they were formed by similarity of closure than by similarity of orientation, supporting a dominance of closure over orientation in guiding perceptual grouping. Quinlan and Wilton [11] examined the tendency of forming perceptual groups defined by different Gestalt laws. They presented subjects with displays comprising a row of seven colored shapes and required them to rate the degree to which the central target shape grouped with either the left or the right flanking shapes. Subjects showed a stronger tendency to group local elements by proximity than by similarity.

* This study was supported by National Natural Science Foundation of China (Project 30225026), the Ministry of Science and Technology of China (Project 2002CCA01000), and Learning & Cognition Lab of Capital Normal University.

Correspondence should be addressed to: Han Shihui, Phone: 62759138; Email: shan@pku.edu.cn
Taken together, the findings indicate that proximity grouping takes place earlier and/or is perceived faster than grouping by similarity of shape in the foveal vision.

This claim is supported by recent neurophysiological studies. For example, Han et al. \(^{12, 13}\) investigated the neural substrates underlying Gestalt grouping by recording high-density event-related brain potentials (ERPs). They asked subjects to discriminate orientations of perceptual groups made up of local elements. They found that proximity grouping was indexed by a positive activity over the medial occipital cortex at 100 ms after sensory stimulation and a subsequent right-occipitotemporal negative activity at 230 ms. Grouping by similarity was reflected only in a long-latency occipitotemporal negative activity peaking at 340 ms. The ERP findings suggest that neural mechanisms for grouping by proximity and similarity are different over time and spatial organization in the brain.

Note that stimulus arrays were usually displayed in the center of the visual field in the aforementioned studies. Thus although the results of these studies demonstrate the difference between proximity and similarity grouping in the foveal vision, they tell little about the differential grouping operations in the peripheral vision. The present study aimed to investigate the difference between proximity and similarity grouping processes in the peripheral vision. We also examined whether grouping by similarity of orientation takes place slower than grouping by similarity of closure in the peripheral vision as that in the foveal vision. Since visual acuity is lower in the peripheral than foveal vision, grouping may not vary as a function of geometrical properties in a local area that define the grouping processes.

2 Experiment 1

Experiment 1 employed a detection task to examine if proximity grouping dominates grouping by similarity of shape in the peripheral vision. Subjects were required to respond to targets in which local arrows or triangles were arranged in rows or columns. The arrows or triangles were either black on a grey background or surrounded by crosses. Local arrows or triangles were grouped based on proximity in the former condition whereas based on similarity of shape in the latter condition. In addition, as arrows and triangles are different from crosses in orientation and closure, respectively, it is possible for us to examine if closure dominates orientation in similarity grouping in the peripheral vision.

2.1 Method

2.1.1 Subjects. Fourteen undergraduate students (5 male, 9 female; aged between 18–23 years) from Peking University participated in Experiment 1 as paid volunteers. All had normal or corrected-to-normal vision.

2.1.2 Apparatus. Data collection and stimulus presentation were controlled by a personal computer. Stimuli were presented on a 17-in. monitor at a viewing distance of about 70 cm.

2.1.3 Stimuli. Stimuli consisted of square arrays of local arrows or triangles, arranged in a 11 × 11 matrix, as illustrated in Figure 1. In half of the stimuli, the local arrows or triangles were black on a grey background and arranged into columns or rows (Figure 1a). In the other half of the stimuli, crosses were inserted between the columns or rows of arrows or triangles (Figure 1b). The size of each cross was the same as that of each local arrow or triangle. The distance between adjacent triangles or arrows was equivalent to that between each triangle or arrow and its neighboring crosses. Proximity dominated grouping of local arrows or triangles when they were presented on a grey background whereas similarity of shapes became dominant when they were displayed with crosses. The global array was 5.42 × 5.42 cm (height and width), and the local figure was 0.40 × 0.40 cm. The global array and each local figure subtended visual angles of 5.16° × 5.16° and 0.38° × 0.38°, respectively. The distance between the fixation and inner edge of the global array was 1.6°. The RGB values, which were kept unchanged, for the background screen were (200, 200, 200) and for the local elements (arrows, triangles, and crosses) were (0, 0, 0).
2.1.4 Procedure. For Task A, each trial began with the presentation of a small cross in the center of the screen as fixation. After 1,000 ms, a stimulus array was presented in the left visual field (LVF) or the right visual field (RVF) for 150 ms. Interstimulus intervals (ISI) were randomized between 850 – 1,150 ms. The stimuli consisted of arrows and triangles presented in separate blocks of trials. There were two blocks of 160 trials for each type of stimuli (arrow vs. triangle). Subjects were required to detect the presence of groups with specific orientations (horizontal or vertical) that were presented in either the LVF or the RVF, regardless of whether proximity or similarity cues dominated grouping. Each block was divided into four epochs. At the beginning of each epoch, a cue was displayed on the screen to indicate orientations of stimuli to respond (horizontal or vertical) until the subject pressed a button to continue. There were 50% targets in each block of trials. Subjects responded by pressing the left button of a mouse with the right index finger. Instructions emphasized both accuracy and speed.

RTs, hits and false alarms were subjected to repeated measures analyses of variance (ANOVAs), with factors being Grouping (proximity vs. similarity), Local Element Shape (arrows vs. triangles), and Visual Field (LVF vs. RVF).

For Task B, all aspects were the same as those in Task A except that two stimulus arrays were presented in the LVF and RVF simultaneously in each trial. Grouping of targets and distractors could be dominated by either proximity or similarity. RTs, hits and false alarms were subjected to ANOVAs, with factors being Target Grouping (proximity vs. similarity), Distractor Grouping (proximity vs. similarity), Local Element
Shape (arrows vs. triangles) and Visual Field (LVF vs. RVF).

For Task C, all aspects were the same as those in Task B except that subjects were required to detect stimulus arrays in the two visual fields that had the same orientation. RTs, hits and false alarms were subjected to ANOVAs, with factors being Target Grouping (proximity in both visual fields, similarity in both visual fields, proximity in one visual field and similarity in the other visual field), Local Element Shape (arrows vs. triangles).

2.2 Results and Discussion

### Table 1 Mean RTs (ms), hit rates (%) and false alarm rates (%) under each condition in Detection Task A of Experiment 1

<table>
<thead>
<tr>
<th>Local Element Shape</th>
<th>Arrow</th>
<th>Similarity</th>
<th>Triangle</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proximity</td>
<td></td>
<td>Proximity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LVF</td>
<td>RVF</td>
<td>LVF</td>
<td>RVF</td>
</tr>
<tr>
<td>RT</td>
<td>408</td>
<td>403</td>
<td>444</td>
<td>431</td>
</tr>
<tr>
<td>Hit</td>
<td>100</td>
<td>99.3</td>
<td>99.3</td>
<td>98.6</td>
</tr>
<tr>
<td>False Alarm</td>
<td>4.1</td>
<td>1.8</td>
<td>2.5</td>
<td>3.9</td>
</tr>
</tbody>
</table>

The mean RTs for correct responses in Task B are shown in Table 2. The main effect of Target Grouping was significant $[F(1, 13) = 142.41, p < 0.0005]$, RTs were faster when detecting proximity than similarity targets (430 vs. 472 ms). The main effect of Distractor Grouping $[F(1, 13) = 8.04, p < 0.05]$ and the interaction between two factors $[F(1, 13) = 47.26, p < 0.0005]$ were also significant, reflecting the fact that the grouping of distractors influenced the detection of targets. Separate analyses showed that the influence of distractors was significant for similarity targets $[F(1, 13) = 31.75, p < 0.0005]$ but not for proximity targets ($p > 0.1$). Responses to similarity targets were faster when distractors shared the same orientation with targets. The effects of Local Element Shape and Visual Field were not significant ($p > 0.1$).

The mean RTs for correct responses in Task A are shown in Table 1. ANOVAs showed a main effect of Grouping $[F(1, 13) = 51.54, p < 0.0005]$, RTs to the proximity – grouping stimuli were faster than those to the similarity – grouping stimuli (408 vs. 440 ms). The effects of Local Element Shape and Visual Field were not significant ($p > 0.1$), nor were interactions between these factors ($F < 1$). The mean hits and false alarms under each condition are shown in Table 1. No main effects or interactions reached significance, indicating that there was no speed – accuracy tradeoff.

The mean RTs for correct responses in Task B are shown in Table 2. The main effect of Target Grouping was significant $[F(1, 13) = 142.41, p < 0.0005]$, RTs were faster when detecting proximity than similarity targets (430 vs. 472 ms). The main effect of Distractor Grouping $[F(1, 13) = 8.04, p < 0.05]$ and the interaction between two factors $[F(1, 13) = 47.26, p < 0.0005]$ were also significant, reflecting the fact that the grouping of distractors influenced the detection of targets. Separate analyses showed that the influence of distractors was significant for similarity targets $[F(1, 13) = 31.75, p < 0.0005]$ but not for proximity targets ($p > 0.1$). Responses to similarity targets were faster when distractors shared the same orientation with targets. The effects of Local Element Shape and Visual Field were not significant ($p > 0.1$).

The mean RTs for Task C are shown in Table 3. ANOVAs showed a reliable main effect of Grouping $[F(3, 39) = 61.27, p < 0.0005]$, RTs to proximity targets in both visual fields were faster than those to similarity targets $[F(1, 13) = 73.32, p < 0.0005]$, which were in turn faster than those to targets that were defined by proximity in one visual field and by similarity in the other visual field $[F(1, 13) = 12.11, p < 0.005]$. The effect of Local Element Shape was also significant $[F(1, 13) = 5.80, p < 0.05]$, responses to the groups made up of local triangles were faster than those to the groups composed of local arrows. However, the interaction between grouping and form was not significant $[F(3, 39) = 1.14, p > 0.3]$.

The mean hits and false alarms under each condition are given in Table 3. The main effects of Grouping on both hits and false alarms were significant ($p < 0.01$), suggesting that response accuracy was higher in perceptual grouping based on proximity than similarity. The effects of Local Element Shape was not significant ($p > 0.1$).
The results of Experiment 1 showed that subjects responded faster in detection of proximity than similarity targets. This was true in all the three tasks in which target stimuli were displayed unilaterally, bilaterally, or accompanied with distractors in the opposite visual field. Basically, these results are consistent with the previous studies\(^{7,10}\). However, they were obtained under the condition when stimulus arrays were presented in the peripheral vision. The findings provide evidence for the difference between proximity and similarity grouping in the peripheral vision. Similar to that in the foveal vision, proximity also dominates similarity of shape in the periphery, suggesting that proximity grouping takes place earlier than similarity grouping in both the foveal and peripheral vision. Note that RTs in Task C were slower under the condition when grouping in the two fields was defined by different Gestalt laws than when grouping was defined by similarity in both visual fields. This apparently conflicts with the model in which proximity grouping occurs faster than similarity grouping. However, subjects made more false alarms in the latter than in the former conditions, suggesting that the RT difference may result from the difference in responses bias.

Unlike the results of the previous studies, we did not find any difference between grouping operations based on similarity of orientation and closure. It is possible that the perception of local element geometric features is necessary for the difference between grouping based on orientation and closure similarity. This was impaired by the lower visual acuity in the peripheral vision and thus led to the disappearance of the difference between orientation and closure similarity grouping in the periphery.

In addition, the difference between proximity and similarity grouping was not distinguishable between the LVF and the RVF, suggesting a hemispheric symmetry for the dominance of proximity over similarity in perceptual grouping.

### 3 Experiment 2

In the detection tasks used in Experiment 1, sub-
jects might make responses based on ambiguous and simple features of perceptual groups. In Experiment 2 we asked to subjects to discriminate orientations (horizontal vs. vertical) of perceptual groups defined by either proximity or similarity. Higher order cognitive processing may be involved in such a discrimination task. We asked whether the difference between proximity and similarity grouping under this condition shows the same pattern as that in detection tasks.

3.1 Method

3.1.1 Subjects. Fourteen undergraduate students (9 male, 5 female; aged between 21 – 23 years) from Peking University participated in Experiment 2 as paid volunteers. All had normal or corrected – to – normal vision.

3.1.2 Apparatus, Stimuli, and Procedure. All aspects were the same as those in Experiment 1 except that three discrimination tasks were employed in Experiment 2.

For Task A, each trial began with the presentation of a small cross in the center of the screen as fixation. After 1,500 ms, a stimulus array was presented in the LVF or the RVF for 150 ms. ISI were randomized between 1,350 – 1,650 ms. The stimuli consisted of arrows and triangles presented in two separate blocks of 160 trials. On each trial, subjects were required to discriminate orientations of the groups (horizontal vs. vertical) composed of arrows or triangles, presented in either the LVF or the RVF, by pressing one of two keys on a standard keyboard with the left and right index fingers regardless of whether proximity or similarity cues produced grouping. The relationship between the orientations of the groups and the responding hand was counterbalanced across subjects.

RTs, hits and false alarms were subjected to ANOVAs, with factors being Grouping (proximity vs. similarity), Local Element Shape (arrows vs. triangles), and Visual Field (LVF vs. RVF).

For Task B, all aspects were the same as those in Task A except for the following. A distractor stimulus consisting of arrows or triangles and crosses (as shown in Figure 2) was displayed with target stimuli but in the opposite visual field. Local arrows or triangles in the distractor stimuli were interleaved with crosses so that no horizontal or vertical groups appeared in the distractors. Subjects were required to discriminate horizontal or vertical orientations of the target groups while ignoring distractor stimuli.

For Task C, all aspects were the same as those in Task A except that two stimulus arrays were presented simultaneously in the LVF and the RVF. Subjects were required to discriminate whether orientations of perceptual groups in the two stimulus arrays were the same or different. RTs, hits and false alarms were subjected to ANOVAs, with factors being Target Grouping (proximity in both visual fields, similarity in both visual fields, proximity in one visual field and similarity in the other visual field), Local Element Shape (arrows vs. triangles).

3.2 Results and Discussion

The mean RTs in Task A are shown in Table 4. ANOVAs showed a significant main effect of Grouping...
...$F(1,13) = 9.88, p < 0.001\right]$), suggesting that subjects made fewer errors in the discrimination of proximity relative to similarity stimuli.

### Table 4 Mean RTs (ms) and accuracies (%) under each condition in Discrimination Task A of Experiment 2

<table>
<thead>
<tr>
<th>Local Element Shape</th>
<th>Target Grouping</th>
<th>Arrow</th>
<th>Visual Field</th>
<th>Proximity</th>
<th>Similarity</th>
<th>Proximity</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LVF</td>
<td>RVF</td>
<td>LVF</td>
<td>RVF</td>
<td>LVF</td>
<td>RVF</td>
</tr>
<tr>
<td>RT</td>
<td></td>
<td>533</td>
<td>529</td>
<td>565</td>
<td>548</td>
<td>531</td>
<td>521</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td>96.4</td>
<td>94.6</td>
<td>92.8</td>
<td>94.8</td>
<td>94.6</td>
<td>94.8</td>
</tr>
</tbody>
</table>

ANOVA of the mean RTs in Task B (see Table 5) showed a reliable main effect of Grouping [$F(1, 13) = 47.25, p < 0.0005\right]$. RTs were faster in discrimination of proximity than similarity stimuli (555 vs. 598 ms). The effects of Local Element Shape and Visual Field were not significant ($p > 0.1\right)$. The mean accuracies rates under each condition are given in Table 5. No main effects or interactions reached significance.

### Table 5 Mean RTs (ms) and accuracies (%) under each condition in Discrimination Task B of Experiment 2

<table>
<thead>
<tr>
<th>Local Element Shape</th>
<th>Target Grouping</th>
<th>Arrow</th>
<th>Visual Field</th>
<th>Proximity</th>
<th>Similarity</th>
<th>Proximity</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>LVF</td>
<td>RVF</td>
<td>LVF</td>
<td>RVF</td>
<td>LVF</td>
<td>RVF</td>
</tr>
<tr>
<td>RT</td>
<td></td>
<td>533</td>
<td>555</td>
<td>602</td>
<td>598</td>
<td>567</td>
<td>546</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td>94.8</td>
<td>95.7</td>
<td>92.9</td>
<td>91.3</td>
<td>92.5</td>
<td>95.5</td>
</tr>
</tbody>
</table>

The mean RTs for correct responses in Task C are shown in Table 6. The main effect of Grouping was significant [$F(3,39) = 31.68, p < 0.0005\right]$. Separate analyses confirmed that RTs to the stimuli grouped by proximity in both LVF and RVF were faster than those to the stimuli grouped by similarity in both hemifields [$F(1,13) = 80.98, p < 0.0005\right]$ and the stimuli in which proximity and similarity dominated grouped respectively in the two hemifields [$F(1,13) = 43.96, p < 0.0005\right]$, however RTs did not differ between the latter two conditions [$F(1, 13) = 1.98, p > 0.1\right]$. The effect of Local Element Shape and its interaction with Grouping were not significant.

### Table 6 Mean RTs (ms) and accuracies (%) under each condition in Discrimination Task C of Experiment 2

<table>
<thead>
<tr>
<th>Local Element Shape</th>
<th>Target Grouping</th>
<th>Arrow</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
</tr>
</thead>
<tbody>
<tr>
<td>RT</td>
<td></td>
<td>593</td>
<td>629</td>
<td>624</td>
<td>624</td>
<td>598</td>
<td>638</td>
<td>634</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td>94.1</td>
<td>90.5</td>
<td>90.5</td>
<td>91.8</td>
<td>85.5</td>
<td>86.6</td>
<td></td>
</tr>
</tbody>
</table>

The mean accuracies under each condition are shown in Table 6. The main effect of Grouping was significant [$F(3,39) = 4.67, p < 0.005\right]$, subjects were more accurate in discriminating proximity than similarity stimuli.

Although discrimination tasks were used in Experiment 2, subjects' performance was quite similar to that in Experiment 1 where detection tasks were used. RTs to proximity stimuli were faster than those to similarity stimuli under the condition when stimulus arrays were presented unilaterally or accompanied with a distractor array in the opposite visual field and when subjects had to discriminate orientations of perceptual groups presented in both hemifields. In addition, our data showed...
that subjects' responses did not differ between the conditions when similarity of orientation and closure determined grouping of local arrows or triangles. Our findings suggest that the difference between proximity and similarity grouping in the peripheral vision is independent of task requirement and thus is determined by the early visual processing.

4 General Discussion

The current study aimed to investigate the difference between grouping defined by proximity and similarity in the peripheral vision. Stimulus arrays composed of local arrows or triangles, which were organized into columns or rows by proximity or similarity of orientation or closure, were presented briefly in the LVF or the RVF. In Experiment 1 subjects were instructed to detect the presence of stimulus arrays in which local arrows or triangles grouped into columns or rows. In Experiment 2 subjects discriminated the orientations (horizontal vs. vertical) of perceptual groups in each trial.

We found that, regardless of whether a detection task or a discrimination task was used, subjects responded faster to proximity than similarity stimuli. The dominance of proximity over similarity in grouping operations was not different between stimuli presented in the LVF or the RVF. These results are in agreement with the previous studies which found a similar pattern in the foveal vision\(^7,\)\(^8\). The findings are consistent with a model in which grouping by proximity takes place earlier than grouping by similarity\(^12,\)\(^13\). Furthermore, the current work suggests that the dominance of proximity over similarity reflects early representation of spatial relationship between local elements and is independent of visual acuity.

Our data showed that the difference between proximity and similarity grouping did not differ between the LVF and the RVF stimuli. Our previous ERP research\(^12,\)\(^13\) found short – latency proximity – grouping related activity over the medial occipital cortex and long – latency enhanced activity over the right hemisphere for proximity grouping but over the left hemisphere for similarity grouping. The behavioral data of the present study may only reflect early dissociation between proximity and similarity grouping in the two hemispheres.

We also found that similarity of orientation is as efficient as similarity of closure in determining perceptual grouping in the periphery. This is different from the results when stimulus arrays are presented foveally\(^8-\)\(^10\). It is possible that grouping by similarity requires identification of local geometrical difference between neighboring local elements. Low visual acuity in the peripheral vision made it difficult to identify or discriminate geometrical properties and thus eliminated the difference between different similarity grouping operations.

References

3 Marr D. Vision. W H Freeman, San Francisco, 1982
5 Koffka K. Principles of gestalt psychology. Harcourt, Brace, New York, 1923
6 Wertheimer M. Untersuchungen zur Lehre von der gestalt; II [Principles of perceptual organization]. Psychologische Forschung, 1923, 4; 301 – 350
7 Ben – A M B, Sagi D. Perceptual grouping by similarity and proximity; Experimental results can be predicted by intensity autocorrelations. Vision Research, 1995, 35; 853 – 866
8 Han S, Humphreys G W. Interactions between perceptual organization based on Gestalt laws and those based on hierarchical processing. Perception & Psychophysics, 1999, 61; 1287 – 1298
12 Han S, Song Y, Ding Y, Yund E W, Woods D L. Neural substrates for visual perceptual grouping in humans. Psychophysiology, 2001, 38(6); 926 – 935
外周视野的相邻性和相似性知觉组织

蒋毅韩世辉

北京大学心理学系
脑与认知科学中心
北京

首都师范大学学习和认知实验室
北京

摘 要 我们先前的实验研究发现在中央视野相邻性原则比相似性原则在知觉组织形成过程中占据主导地位。本研究采用靶目标探测任务和分辨任务进一步检验这两种组织原则在外周视野的差异。实验要求被试判断由基于相邻性或相似性形成的知觉组织的朝向，刺激呈现在左侧或者右侧视野。实验结果表明在两种任务条件下被试对相邻性组织的反应时间都快于相似性组织，相邻性对相似性的这种优势在左右视野没有差异，并且由朝向和封闭性相似性决定的知觉组织也没有差别。这样的实验结果提示在外周视野由相邻性决定的知觉组织比相似性决定的知觉组织发生的要早，而且由不同几何性质决定的相似性组织的加工速度相同。

关键词 知觉组织 相邻性 相似性 外周视野

分类号